Trường đại học Bách Khoa Hà Nội

Viện Công Nghệ Thông Tin và Truyền Thông

──────── \* ───────



HỆ CƠ SỞ DỮ LIỆU ĐA PHƯƠNG TIỆN

**Đề tài: Ứng dụng phân đoạn video**

*Giảng viên hướng dẫn*: **TS. Nguyễn Thị Oanh**

*Sinh viên thực hiện*: Nguyễn Văn Tuân – 20144838

Đào Duy Hòa – 20151589

Vũ Đức Anh Tuấn – 20154153

***Hà Nội, 12/2018***

1

[I. Mô tả bài toán 2](#__RefHeading___Toc1002_3721583951)

[II. Phương pháp xây dựng hệ thống 3](#__RefHeading___Toc1004_3721583951)

[2.1 Cách tính điểm đặc trưng 3](#__RefHeading___Toc1006_3721583951)

[2.2 Cách tính độ tương đồng 3](#__RefHeading___Toc1008_3721583951)

[2.3 Cách xác định ngưỡng 4](#__RefHeading___Toc1010_3721583951)

[2.4 Cách xác định keyframe 4](#__RefHeading___Toc1012_3721583951)

[2.5 Phương pháp triển khai 4](#__RefHeading___Toc1014_3721583951)

[III. Ưu/ nhược điểm 5](#__RefHeading___Toc1016_3721583951)

[IV. Thực nghiệm 6](#__RefHeading___Toc1018_3721583951)

[IV. Đánh giá hiệu năng của hệ thống 7](#__RefHeading___Toc1020_3721583951)

## **I. Mô tả bài toán**

Đề tài : Phân đoạn video đơn giản (phát hiện các shot) sử dụng đặc trưng cục bộ.

Một video là một phép toán kết hợp bởi nhiều shot- đơn vị nhỏ nhất tạo nên video. Mỗi shot là một cảnh quay liên tục được tạo bởi 1 camera.

Mỗi shot cấu thành bởi nhiều frame, tỉ lệ điểm 2 frame liên tiếp trong cùng 1 shot có độ tương đồng tương tự nhau.

Độ tương đồng được đo bằng số điểm đặc trưng được so khớp trên tổng số điểm được phát hiện.

## **II. Phương pháp xây dựng hệ thống**

### **2.1 Cách tính điểm đặc trưng**

Để giải quyết bài toán trên, nhóm đã sử dụng **Sift** để tìm các điểm đặc trưng cục bộ rồi tính tỉ lệ độ tương đồng. Dựa vào đó chọn ra ngưỡng để phân đoạn video.

Để phân biệt **keypoint** này với keypoint khác cần tìm ra tham số đại diện gọi là **descriptor**. Hai keypoint khác nhau thì descriptor khác nhau.

Sử dụng **detectAndCompute()** để tính ra các điểm đặc trưng **keypoint** và các **descriptor** tương ứng.

**kp1, des1 = sift.detectAndCompute(gray1, None)**

### **2.2 Cách tính độ tương đồng**

Bước 1 : Tính điểm đặc trưng của từng ảnh sử dụng detectAndCompute()

**kp1, des1 = sift.detectAndCompute(gray1, None)**

**kp2, des2 = sift.detectAndCompute(gray2, None)**

Bước 2 : Tính điểm đặc trưng khớp nhau giữa 2 frame dùng BFMatcher()

**bf = cv.BFMatcher()**

**matches = bf.match(des1, des2)**

Bước 3 : Tính độ tương đồng giữa 2 frame

**ratio =**

### **2.3 Cách xác định ngưỡng**

Nhận thấy tại biên của các shot xảy ra sự chênh lệch lớn về độ tương đồng, giá trị này trội lên nhiều so với các giá trị khác. Do đó, nhóm sẽ lấy giá trị này để làm ngưỡng để phân đoạn video.

Dựa vào tỉ lệ độ tương đồng giữa các frame được vẽ trên biểu đồ xác định 1 ngưỡng thích hợp để sử dụng làm ngưỡng.

### **2.4 Cách xác định keyframe**

Mỗi shot lấy ra một frame đại diện gọi là keyframe. Keyframe là khung hình chính có sự khác biệt với các frame còn lại.

Trong báo cáo này, nhóm chọn keyframe là frame có số điểm đặc trưng cao nhất trong shot.

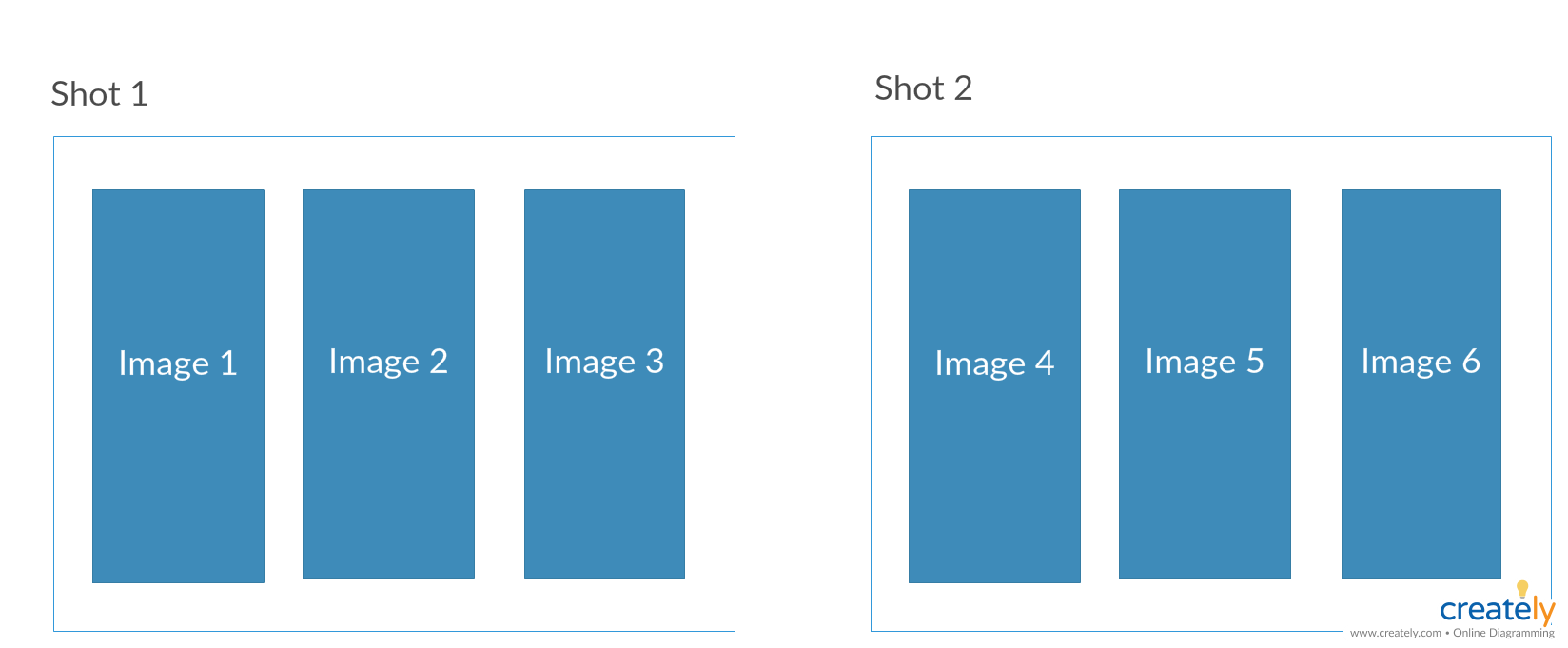
### **2.5 Phương pháp triển khai**

Bước 1 : Đọc, tách video ra thành các frame riêng biệt.

Bước 2 : Tính các điểm đặc trưng từng frame.

Bước 3 : Tính tỉ lệ độ tương đồng 2 frame liên tiếp ( frame 1, frame 2 )

Bước 4 :Tính tỉ lệ độ tương đồng 2 frame tiếp theo ( frame 2, frame 3 )



Bước 5 : Tính giá trị tuyệt đối hiệu của 2 tỉ lệ trên.

Bước 6 : Dựa vào biểu đồ tỉ lệ ở bước 5 => xác định ngưỡng thích hợp.

Bước 7 : Kiểm tra tỉ lệ nào lớn hơn ngưỡng thì phân đoạn ở đó.

## **III. Ưu/ nhược điểm**

Ưu điểm :

- Đưa ra kết quả ổn định với những scale ảnh khác nhau.

- Sử dụng được nội dung trong từng frame nên chính xác hơn phương pháp sử dụng đặc trưng toàn cục.

Nhược điểm :

- Chọn ngưỡng tùy vào từng loại video.

- Chọn ngưỡng quá thấp : false cuts ( phát hiện sai )

- Chọn ngưỡng quá cao : missed cuts ( phát hiện thiếu )

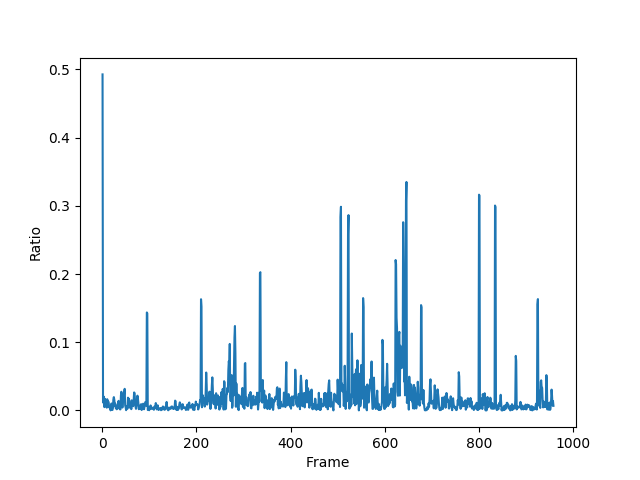
- Tốc độ trích xuất keypoint và chi phí đối sánh lớn với những video có kích thước lớn : hd, full hd, 4k.

## **IV. Thực nghiệm**

Video thực nghiệm từ giây thứ 5 => 45 :

<https://www.youtube.com/watch?v=D86RtevtfrA&t>

Biểu đồ tỉ lệ



Một số keyframe của các shot cắt được từ video trên



## **IV. Đánh giá hiệu năng của hệ thống**

Hiệu năng độ chính xác của hệ thống phụ thuộc vào giá trị ngưỡng được chọn để phân đoạn video.

Phụ thuộc chất lượng hình ảnh của các frame trong video : Chất lượng frame càng tốt thì độ tương đồng thu được càng tốt tuy nhiên chi phí đối sánh sẽ tăng theo.